

# Suurandmed ja AI: globaalsed trendid

# R kioskgo

## R kiosk go

VÄIKE RÕÖM

KÜLM AMPS

VÄRSKE TOIT & JOOK

KUUM JOOK & SAIAKE



**VAIKESED  
ANDMED ON KA  
OLULISED**

**Kui suur on suur?  
Kui väike on väike?**



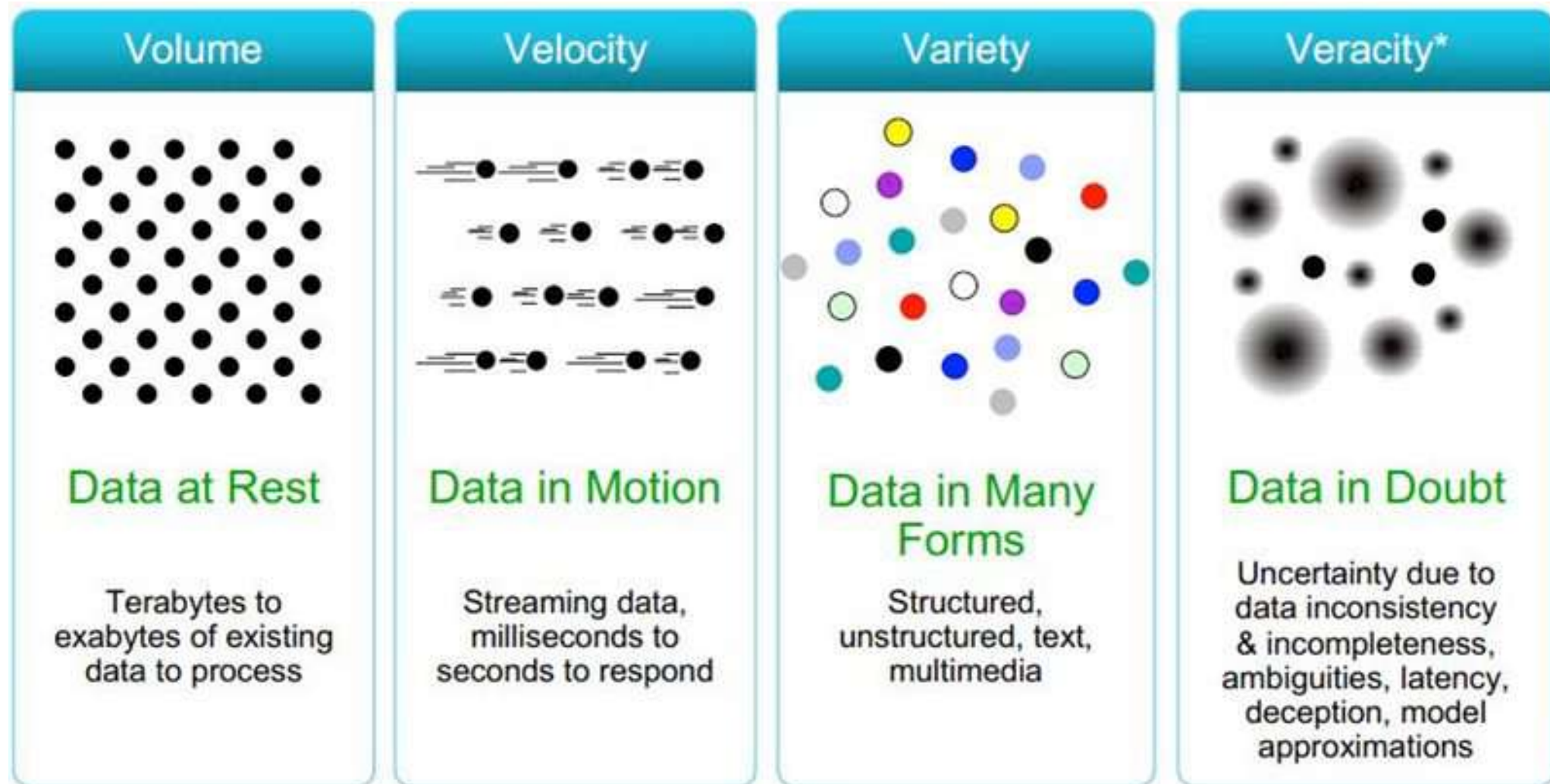
3000 000 MB

3 MB

0.03MB



# Suurandmed: nelja V mudel



# Trendid andmete analüüsimisel

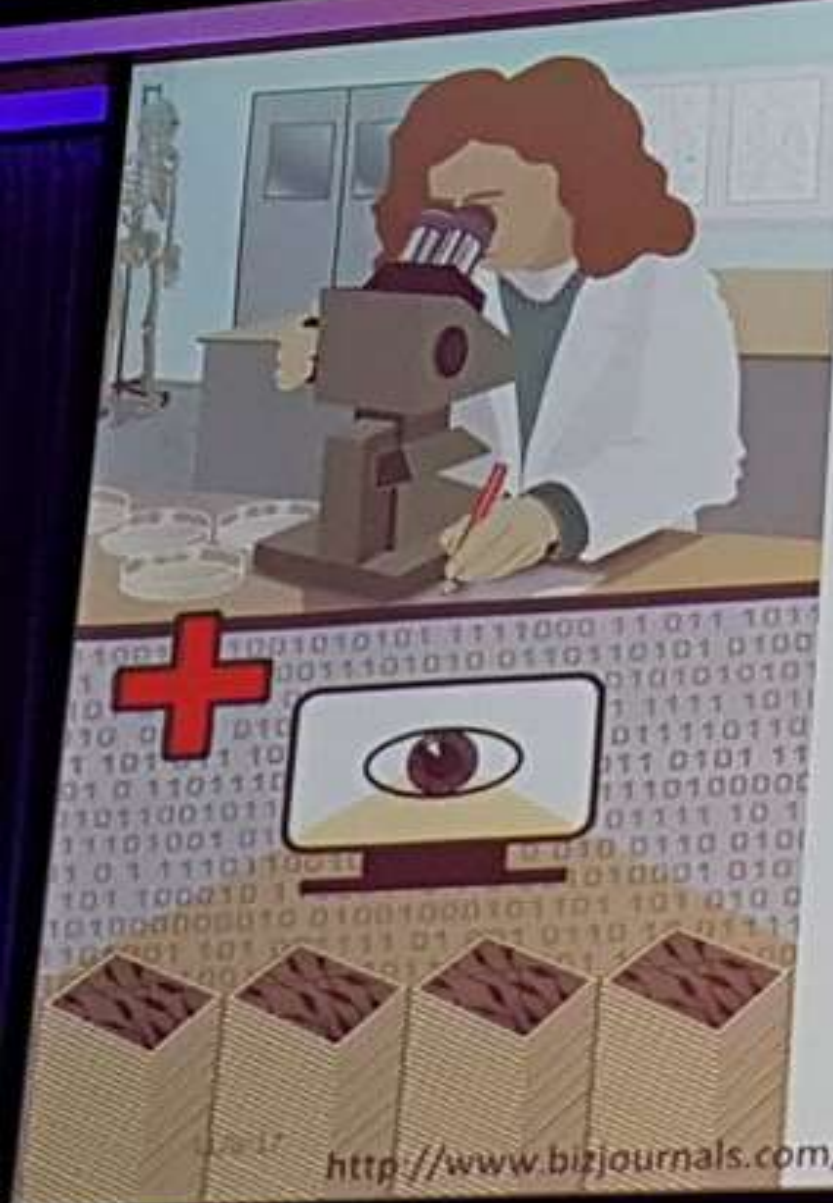
- Social Data Revolution
- Andreas Weigend,
  - ex-Chief Scientist @ Amazon.com
  - Täna lector @ Stanford, UC Berkeley, Tsinghua
- Kaks andmete tekke “revolutsiooni”
  - Klassikaliselt: kaalutletud kogumine
  - 1. laine: igal pool tekib palju, *secondary data*, mõtestamine
  - 2. laine: inimesed ise panevad motivatsiooni korral infot enda hetkeseisundi ja eelistuste kohta (staatused, twitter, linkedin, google)

# Cancer Diagnose Error

Human 3.5%

AI 7.5%

AI + Human 0.5%



<https://hms.harvard.edu/news/better-together>

© 2017 Daniela Rus, CSAIL MIT  
<http://www.bizjournals.com/boston/blog/bioflash/2016/06/beth-israel-researchers-say-ai-can-lower-error.html>

# **TREND 2: UUDSED ANDMEALLIKAD**

- MOBIILI ASUKOHAANDMED, OSTUAJALUGU
- INTERNETIS OLEVAD ANDMED (TWITTER, DELFI KOMMENTAARIUM)
- TEKST KUI ANDMED. SEADUSED KUI ANDMED.
- SATELLIIDI PILDIST PERSONAALSETE NUTIKELLADENI
- VIDEO JA PILT KUI ANDMED. HELI KUI ANDMED.

**FOOKUS: TUVASTAMINE JA SIDUMINE**



# Trend 3: Pildituvastus



4.

Tulevikutrend on  
andmete-põhise  
innovatsiooniga inimesi  
ja ettevõtteid mõõta ja  
suunata





I LOST  
MEDICINE  
My place =  
#humantrafficking

ALGORITHM  
ELITISM

OF QUAL  
GETS A  
!!!

ALGORITHM  
LOST MY  
FUTURE  
PISSED OFF

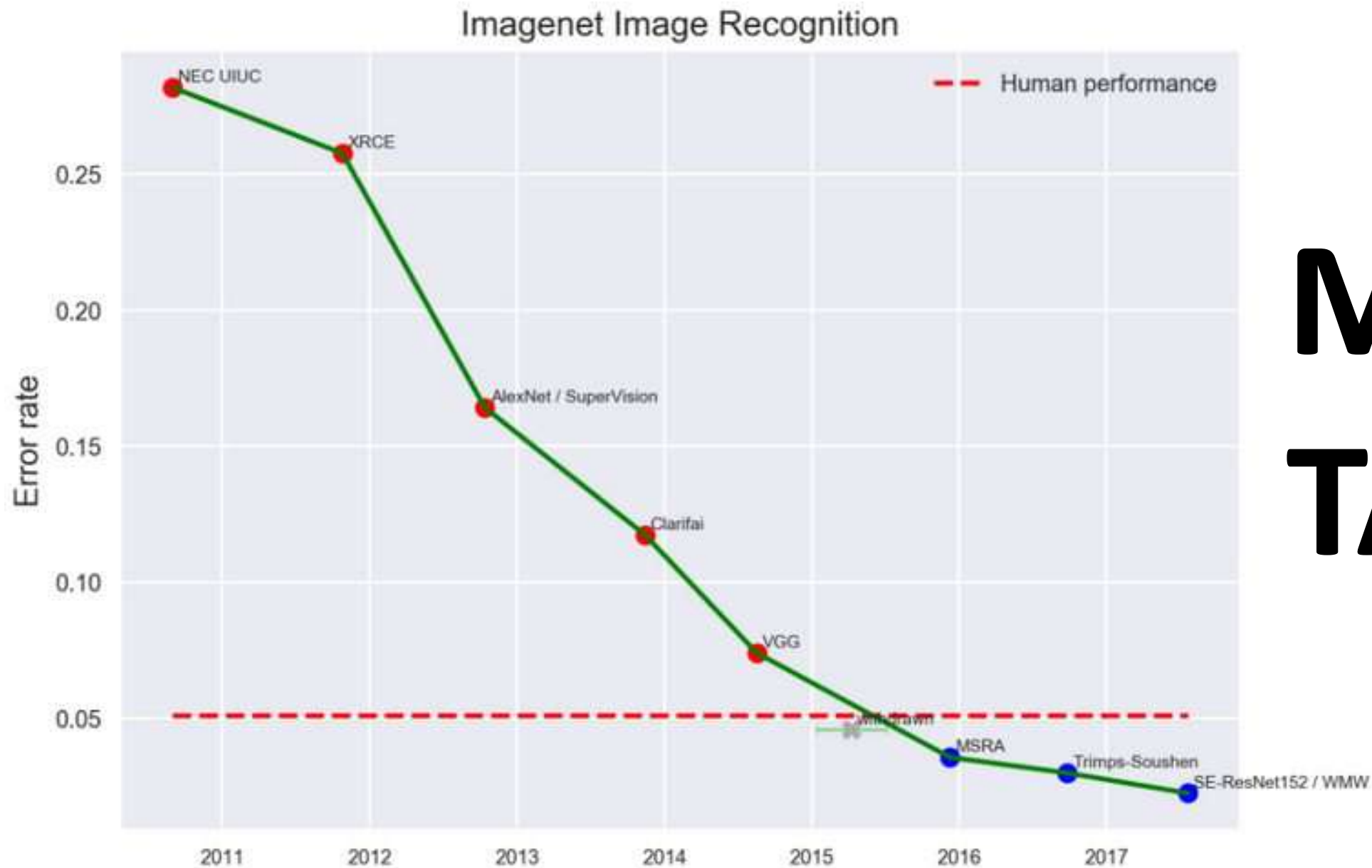
You have  
to  
give us the  
DE

COMPSCI  
STUDENTS COULD  
HAVE MADE A  
BETTER ALGORITHM

DENTS  
OT  
ATS





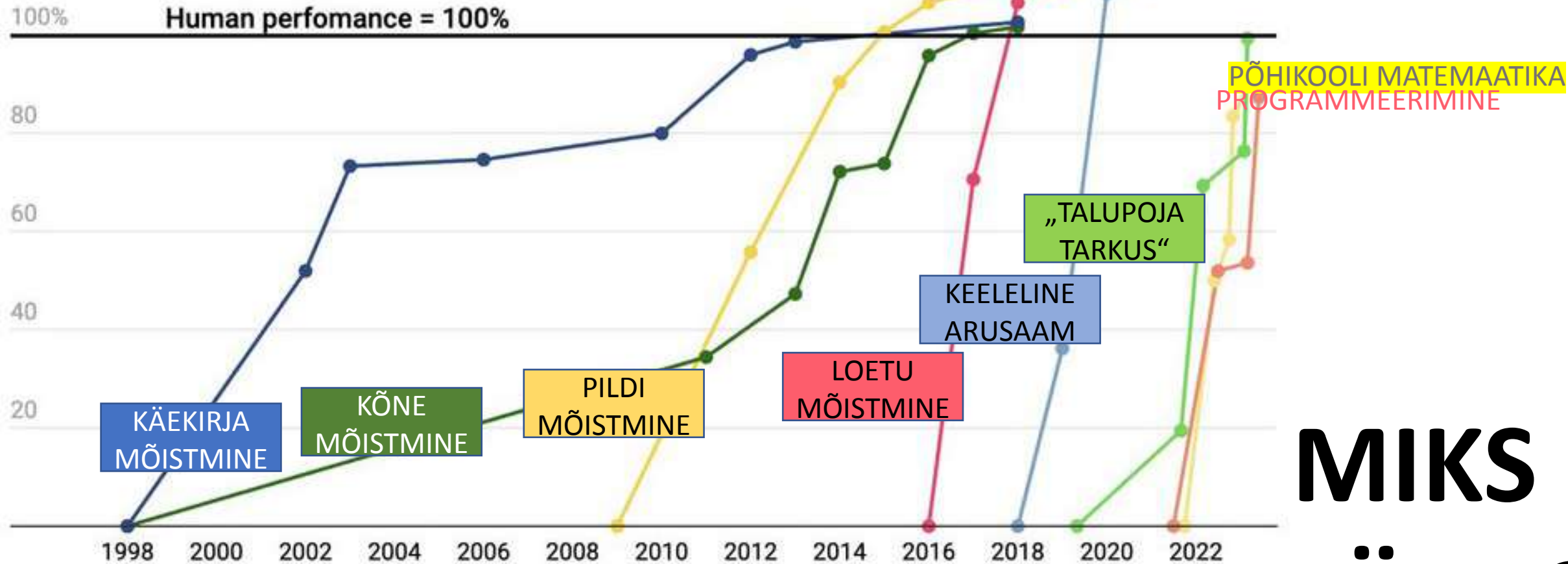


# MIKS TÄNA?

MIS SEIS ON?

INIMVÕIMEKUS = 100%

Human performance = 100%



# MIKS TÄNA?

For each benchmark, the maximally performing baseline reported in the benchmark paper is taken as the "starting point", which is set at 0%. Human performance number is set at 100%. Handwriting recognition = MNIST, Language understanding = GLUE, Image recognition = ImageNet, Reading comprehension = SQuAD 1.1, Reading comprehension = SQuAD 2.0, Speech recognition = Switchboard, Grade school math = GSK8k, Common sense completion = HellaSwag, Code generation = HumanEval.

Chart: Will Henshall for TIME • Source: [ContextualAI](#)

TIME

# November 30: Uus tasuta töötaja ChatGPT

- Keskpäraselt või nõrgalt sooritanud ära:

👨‍⚕️ United States Medical Licensing Exam

🎓 MBA-level Operations exam

👨‍⚖️ The **Bar** Exam (based on typical exam questions)

+ TalTech andmebaaside aine hindele 3 ja lugu Sirpi.

Proovi järele!



# HIRM: ROBOTID TULEVAD JA VÕTAVAD MEIE TÖÖ ÄRA (JA TAPAVAD MEID ÄRA)

- Juhti ei asenda AI, vaid konkurent, kes kasutab AI-d
- AI vajab juhti ja paremaid andmeid
- Kelle asi? „Keda kotib”? Ei ole IT-spetsi, vaid juhi asi
- Poliitika: mikrosihtimisest 1:1 veenmiseni

A man with short brown hair and glasses, wearing a dark suit, light blue shirt, and striped tie, is speaking at a podium. He is looking slightly to his right. The background is blurred, showing what appears to be a conference or event setting.

**AI EI OLE KÕIKVÕIMAS**

**AI VAJAB JUHTI, PAREMAID ANDMEID**

**AI ON TUNDLIKUM KUI INIMENE,  
INIMENE JA MEDITSIIN MUUTUB**

**OHT: ARSTI AEGA KULUB VEEL ROHKEM  
VÕIMALUS: KINDLAD ÜLESANDED**

# (Tehnoloogia ja poliit)**filosoofia** **nurgake**

- Kuidas mõõta (suurandmetega) inimese kannatusi?
- Kui me suudaks mõõta kannatusi, kus me seda kasutada saaks?
- Karistuspoliitika liigub ennetamise suunas, kas kunagi ka tervishoiupoliitika?  
**Kuidas saaks tervisekäitumist arvestada üksikisiku (sotsiaal?)maksukoormuse määramisel?**



# VÄLJAKUTSE: DEANONÜMISEERIMINE

- Tegevus, kus anonüümised andmed riskasutatakse teiste allikatega, et anonüümset allikat reidentifitseerida.

Kuulsad näited: AOL, NETFLIX

 [www.telegraph.co.uk/technology/google/10769786/Britons-trust-Google-more-than-their-families](http://www.telegraph.co.uk/technology/google/10769786/Britons-trust-Google-more-than-their-families)

[HOME](#) » [TECHNOLOGY](#) » [GOOGLE](#)

## Britons trust Google more than their families

The British nation is more inclined to trust Google over their own families, new research reveals

# DEANONÜMISEERIMINE: NÄIDE

AVALIK INFO:  
KÕIGI TÖÖTAJATE PALK  
KOKKU = **EUR 500 000**  
INIMESTE ARV = 8

AVALIKUD ANDMED

NIMI	PALK
INIMENE 1	<u>EI ÜTLE! SEE ON ISIKLIK!!!</u>
INIMENE 2	60 000
INIMENE 3	60 000
INIMENE 4	60 000
INIMENE 5	60 000
INIMENE 6	60 000
INIMENE 7	60 000
INIMENE 8	60 000

# NYC TAXI OPEN DATA CASE



**FOIL REQUEST FORM**

Taxi and Limousine Commission  
Office of Legal Affairs  
33 Beaver Street, 22<sup>nd</sup> Floor  
New York, New York 10004  
Attn: Records Access Officer

FROM: (Please print your name, address, telephone # and email address)  
Mr. Christopher M. Whong  
[REDACTED]  
New York, NY

Our email address: FOIL@tlc.nyc.gov

Your email address: [REDACTED]

I request the following record(s) under the Freedom of Information Law:  
Please reasonably describe the record(s) you are requesting to allow us to identify the document(s) required.

NYC taxi trip data in machine readable format from 1/1/2013 through  
the most current available date.

Please state the reason for your request: (optional)  
Data Analysis

[REDACTED]  
Signature

3/11/2014

Date

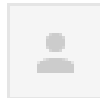
This is a Freedom of Information Law ("FOIL") request. As such, your request will be considered under the Public Officer's Law, Article 6, Section 84 et seq. Subject to the provisions of this article, the Taxi and Limousine Commission, within five business days of the receipt of a written request for a record reasonably described, shall furnish a written acknowledgement of the receipt of such request.



## RE: FOIL Request # 20203



Inbox x



**TLCmail2, TLCmail (TLC)**

Mar 11 (7 days ago) ☆



to me ▾

Mr. Whong,

Pursuant to FOIL, your request for GPS data has been granted. We currently have 2013 data available from January to December. Please be advised in order to obtain this information you will need to send or bring an external hard drive with a minimum capacity of 200 GB to the TLC offices. The address is listed below. The hard drive must be brand new, still in the box and unopened. Should you decide to mail the hard drive please be sure to include return shipping. If you decide to come in person, please let me know when and I will determine my availability.

If you have further questions pertaining to your request please feel free to call our office at the number below and please refer to your FOIL #20203.

Thank You,



Taxi and Limousine Commission (TLC)

Legal Affairs

33 Beaver Street, 22nd Floor

New York, NY 10004

tel: [212-676-1135](tel:212-676-1135)

email: [foil@tlc.nyc.gov](mailto:foil@tlc.nyc.gov)



Chris Whong <

Mar 17 (1 day ago) ☆



to TLCmail ▾

Greetings,

I would like to come in person to retrieve this data. Do you have availability this afternoon (Monday, 3/17)? Will I be able to get the data in one visit or will I need to make arrangements to pick it up?

Also, can you answer the following questions?

- 1) Why do you require the hard disk I provide to be new/unopened?
- 2) When will data after 12/31/2013 become available?

Thank you,

\*\*\*



TLCmail2, TLCmail (TLC)

Mar 17 (1 day ago) ☆



to me ▾

Mr. Whong,

You can come in this afternoon to drop off the hard drive and make the arrangements to retrieve it once completed. The data transfer from our system usually takes a couple hours so we advise you to come back another day and pick it up. We require a brand new unopened hard drive to ensure the protection of our systems from viruses. Our 2014 GPS data from January to June will be available in July 2014.

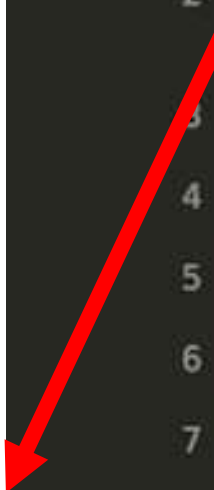
Thank you,

\*\*\*

Taxi and Limousine Commission (TLC)  
Legal Affairs

FOIL #20203	
Name	Size
▼ FOIL2013	--
▼ Faredata_2013	--
trip_fare_1.csv	1.68 GB
trip_fare_2.csv	1.59 GB
trip_fare_3.csv	1.79 GB
trip_fare_4.csv	1.72 GB
trip_fare_5.csv	1.74 GB
trip_fare_6.csv	1.64 GB
trip_fare_7.csv	1.58 GB
trip_fare_8.csv	1.44 GB
trip_fare_9.csv	1.61 GB
trip_fare_10.csv	1.71 GB
trip_fare_11.csv	1.64 GB
trip_fare_12.csv	1.59 GB
▼ Tripdata_2013	--
trip_data_1.csv	2.46 GB
trip_data_2.csv	2.33 GB
trip_data_3.csv	2.62 GB
trip_data_4.csv	2.52 GB
trip_data_5.csv	2.55 GB
trip_data_6.csv	2.4 GB
trip_data_7.csv	2.3 GB
trip_data_8.csv	2.1 GB
trip_data_9.csv	2.35 GB
trip_data_10.csv	2.5 GB
trip_data_11.csv	2.4 GB
trip_data_12.csv	2.33 GB

```
trip_data_1.csv
1 medallion, hack_license, vendor_id, rate_code, store_and_fwd_flag,
  pickup_datetime, dropoff_datetime, passenger_count, trip_time_in_secs,
  trip_distance, pickup_longitude, pickup_latitude, dropoff_longitude,
  dropoff_latitude
2 89D227B655E5C82AECF13C3F540D4CF4, BA96DE419E711691B9445D6A6307C170, CMT, 1, N, 2013-01-
  01 15:11:48, 2013-01-01 15:18:10, 4, 382, 1.
  00, -73.978165, 40.757977, -73.989838, 40.751171
3 0BD7C8F5BA12B88E0B67BED28BEA73D8, 9FD8F69F0804BDB5549F40E9DA1BE472, CMT, 1, N, 2013-01-
  06 00:18:35, 2013-01-06 00:22:54, 1, 259, 1.50, -74.006683, 40.731781, -73.994499, 40.75066
4 0BD7C8F5BA12B88E0B67BED28BEA73D8, 9FD8F69F0804BDB5549F40E9DA1BE472, CMT, 1, N, 2013-01-
  05 18:49:41, 2013-01-05 18:54:23, 1, 282, 1.10, -74.004707, 40.73777, -74.009834, 40.726002
5 DFD2202EE08F7A8DC9A57B02ACB81FE2, 51EE87E3205C985EF8431D850C786310, CMT, 1, N, 2013-01-
  07 23:54:15, 2013-01-07 23:58:20, 2, 244, .70, -73.974602, 40.759945, -73.984734, 40.759388
6 DFD2202EE08F7A8DC9A57B02ACB81FE2, 51EE87E3205C985EF8431D850C786310, CMT, 1, N, 2013-01-
  07 23:25:03, 2013-01-07 23:34:24, 1, 560, 2.10, -73.97625, 40.748528, -74.002586, 40.747868
7 20D9ECB2CA0767CF7A01564DF2844A3E, 598CCE5B9C1918568DEE71F43CF26CD2, CMT, 1, N, 2013-01-
  07 15:27:48, 2013-01-07
  15:38:37, 1.648, 1.70, -73.966743, 40.764252, -73.983322, 40.743763
```



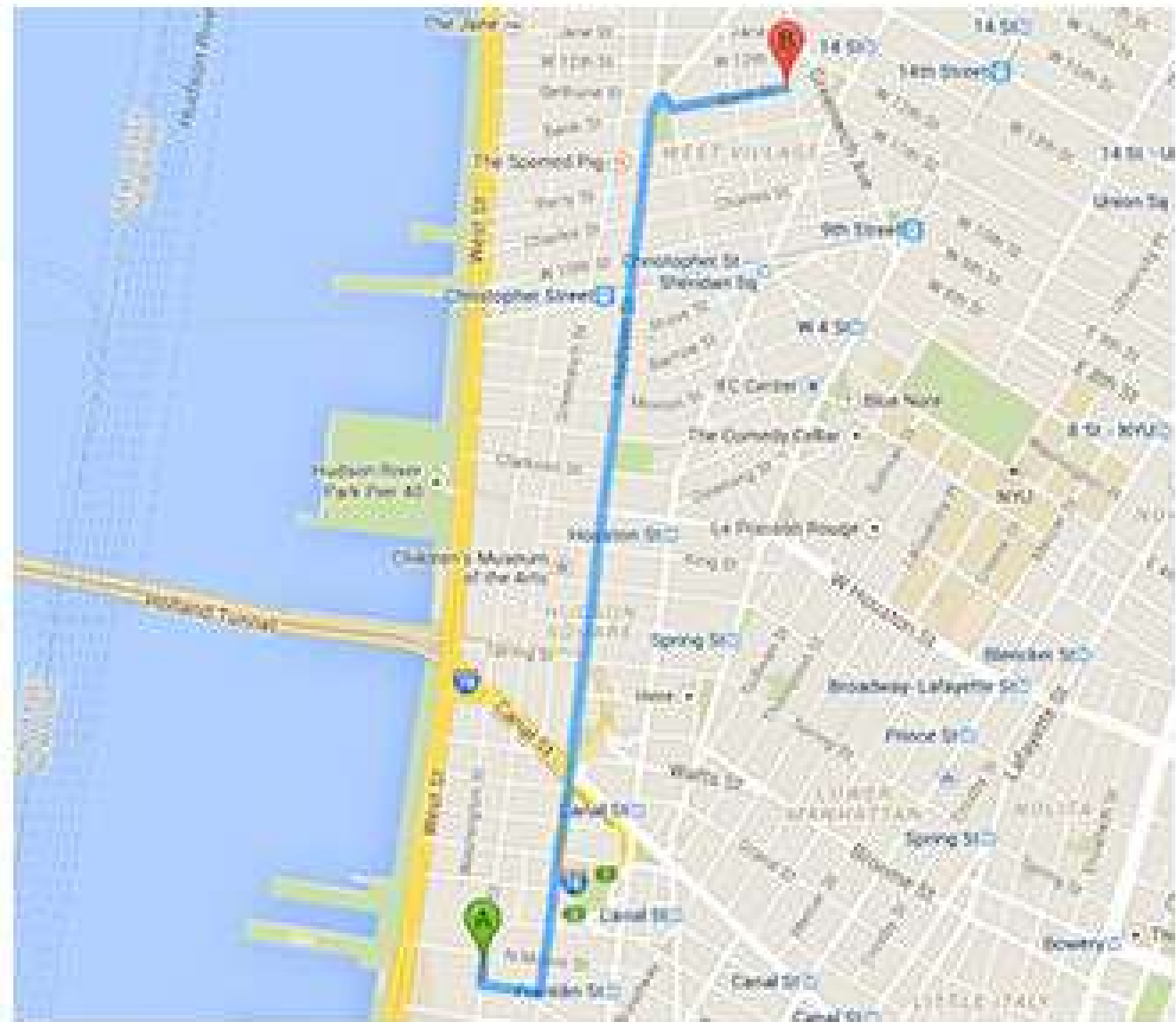
medallion, hack\_license, vendor\_id, rate\_code, store\_and\_fwd\_flag,  
pickup\_datetime, dropoff\_datetime, passenger\_count, trip\_time\_in\_secs,  
trip\_distance, pickup\_longitude, pickup\_latitude, dropoff\_longitude,  
dropoff\_latitude  
89D227B655E5C82AECF13C3F540D4CF4, BA96DE419E711691B9445D6A6307C170, CMT,  
01 15:11:48, 2013-01-01 15:18:10, 4, 382, 1.  
00, -73.978165, 40.757977, -73.989838, 40.751171

```
13 E12F6AF991172EAC3553144A0AF75A19, 06918214E951FA0003D1CC54955C2AB0, CMT, 1, N, 2013-01-
  08 09:50:05, 2013-01-08 10:02:54, 1, 768, .70, -73.99556, 40.749294, -73.988686, 40.759052
14 78FFD9CD0CDA541F335EF8B38FB494D6, E949C583ECF62C8F03FDCE1484954A08, CMT, 1, N, 2013-
```





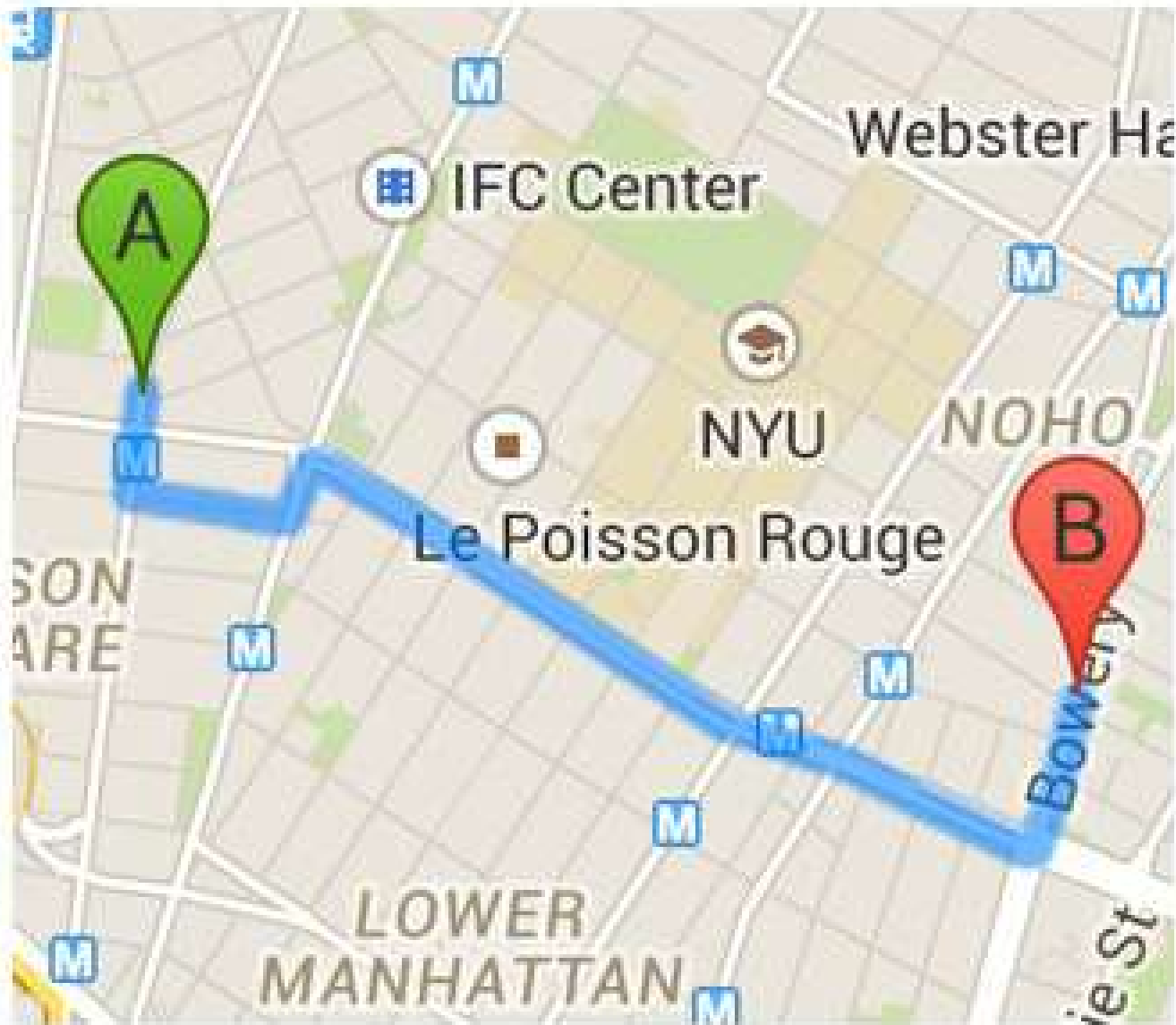
BRADLEY COOPER



JULY 8, 2013 • 7:34 PM - 7:44 PM  
376 GREENWICH ST. TO 13 BANK ST.  
\$9.00 FARE • CASH; UNKNOWN TIP • ©SPLASH



OLIVIA MUNN



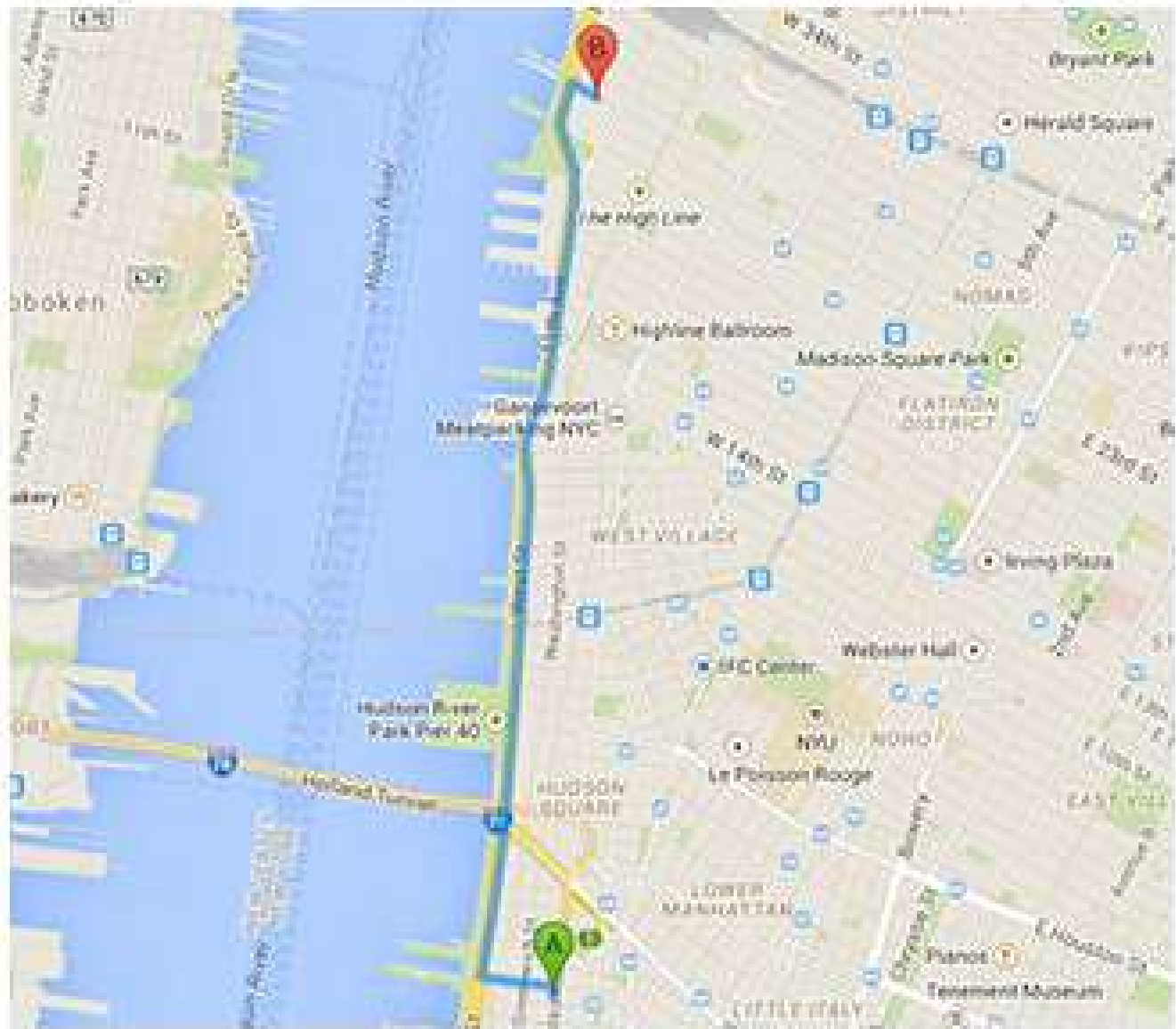
JULY 8, 2013 • 11:20 AM - 11:26 AM

225 VARICK ST. TO 325 BOWERY

\$6.00 FARE • CASH; UNKNOWN TIP • ©SPLASH



KAROLÍNA KURKOVÁ



JUNE 26, 2013 • 11:23 AM - 11:32 AM  
112 HUDSON ST. TO 624 W. 26TH ST.  
\$10.00 FARE • \$2.00 TIP • ©SPLASH





Source: <https://research.neustar.biz/2014/09/15/riding-with-the-stars-passenger-privacy-in-the-nyc-taxicab-dataset/>



🕒 11:17 am




 Legend

- Full Taxi
- Empty Taxi
- Trip Start Point
- Trip End Point
- Trip Route

 Load Another Tax!

© Running Totals

Surcharge: \$3.00

**Tips:** \$19.90

<b>Total:</b>	<b>\$257.73</b>
---------------	-----------------

Passengers: 15





# How anonymous are you?

# 2023

Pause



00:01 / 01:52

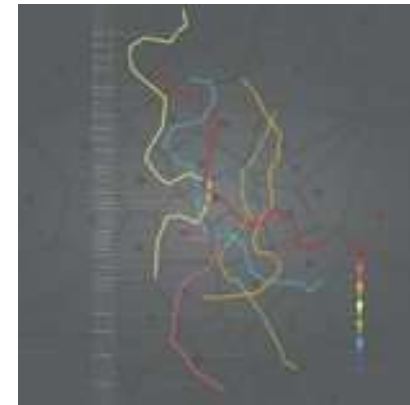


# No such thing as anonymous data

<http://ther.org/2015/02/theres-no-such-thing-as-anonymous-data>



- MIT studied 3 months of **credit card records** for 1.1 million people and show that four spatiotemporal points are enough to uniquely reidentify 90% of individuals. Knowing the price of a transaction increases the risk of reidentification by 22%, on average.
- MIT used 15 months of **telco data** from 1.5 million people to show that 4 points--approximate places and times--are enough to identify 95% of individuals in a mobility database.







**Егор Цветков**

21 год

# YOUR FACE IS BIG DATA





















<https://qz.com/1069756/using-artificial-intelligence-to-identify-protestors-wearing-hats-or-scarves-is-entirely-possible/>



# Artificial Intelligence Frontier: Image Recognition







G



Please tell me what is it in each cell in a tabular format aligned with this picture grid.



Certainly! Here's a breakdown of the content of each cell in the format of a table:



Row	Column 1	Column 2	Column 3	Column 4
1	Muffin with blueberries	Chihuahua (side profile)	Blueberry Muffin	Chihuahua (profile looking up)
2	Chihuahua (frontal face)	Cookie with chocolate chips	Chihuahua (face close-up)	Blueberry Muffin with sugar on top
3	Blueberry Muffin with a smiley face	Chihuahua (frontal face close-up)	Blueberry Muffin with blueberries on top	Chihuahua (slightly showing teeth)
4	Chihuahua (puppy face close-up)	Group of Cookies with chocolate chips	Group of Chihuahua puppies	Muffin with blueberries

Automaatne vestlus

Churn.xls...Innar Liiv

FailAvalahtLisaKõigendusedValerimidAndmedLõbivaatusVaadeJuhata m

Kleebi

Arial10

JoondusArv

TingimusvormingVorminda tabelinaLahtrilaadid

LahtridRedigeerimine

Lõikelaud

Font

Laadid

A1Account Length

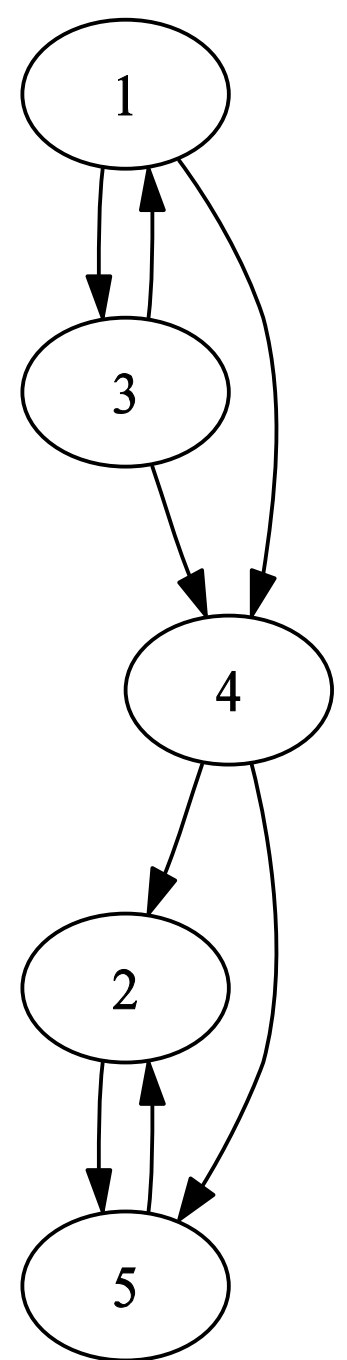
	A	B	C	D	E	F	G	H	I	J	K	L
	Account Length	Mail Message tag	Mins ve	Mins ight	Mins st	Serv Calls	Churn	M'T Plan	Mail Plan tag	Calls tag	Chai	
2	328	25	265.1	197.4	244.7	90	1	0	0	1	110	41
3	907	26	1616	195.5	254.4	12.7	1	0	0	1	122	2
4	137	0	242.4	121.2	162.6	12.2	0	0	0	0	114	4
5	84	0	299.4	61.9	196.9	6.6	2	0	1	0	71	1
6	75	0	166.7	148.3	186.9	10.1	3	0	1	0	113	21
7	118	0	223.4	220.6	203.9	6.3	0	0	1	0	98	3
8	121	24	218.2	348.5	212.6	7.5	3	0	0	1	88	3
9	147	0	157	103.1	211.8	7.1	0	0	1	0	79	21
10	117	0	184.5	351.6	215.8	8.7	1	0	0	0	97	3
11	141	37	258.6	222	326.4	11.2	0	0	1	1	84	4
12	65	0	129.1	228.5	208.8	12.7	4	1	0	0	137	2
13	74	0	197.7	163.4	186	9.1	0	0	0	0	127	2
14	168	0	128.8	104.9	141.1	11.2	1	0	0	0	96	1
15	95	0	156.6	247.6	192.3	12.3	3	0	0	0	88	24
16	62	0	120.7	307.2	203	12.1	4	0	0	0	70	24
17	161	0	332.9	317.8	160.6	5.4	4	1	0	0	67	54
18	85	27	196.4	280.9	83.3	13.8	1	0	0	1	139	3
19	93	0	190.7	288.2	129.6	8.1	3	0	0	0	114	31
20	76	33	189.7	212.8	165.7	90	1	0	0	1	66	31
21	73	0	224.4	159.5	192.8	13	1	0	0	0	90	3
22	147	0	155.1	239.7	208.8	10.6	0	0	0	0	117	24
23	77	0	62.4	169.9	209.6	5.7	5	1	0	0	89	1
24	130	0	183	72.9	191.8	9.5	0	0	0	0	112	1
25	111	0	110.4	137.3	189.6	7.7	2	0	0	0	103	11
26	132	0	81.1	245.2	237	10.3	0	0	0	0	86	1
27	174	0	124.3	277.1	250.7	15.5	3	0	0	0	76	1
28	57	39	213	191.1	182.7	9.5	0	0	0	1	115	3
29	54	0	134.3	155.5	102.1	14.7	3	0	0	0	73	21
30	20	0	190	258.2	181.5	6.3	0	0	0	0	109	1
31	49	0	119.3	215.1	179.7	11.1	1	0	0	0	117	21
32	142	0	84.8	136.7	250.5	14.2	2	0	0	0	95	1
33	75	0	226.1	201.5	246.2	10.3	1	0	0	0	105	31
34	172	0	212	31.2	293.3	12.6	3	0	0	0	121	34
35	12	0	249.6	252.4	280.2	11.8	1	1	0	0	118	4
36	57	25	176.8	195	213.5	8.3	0	0	0	1	94	34
37	72	37	220	217.3	152.8	14.7	3	0	0	1	80	1
38	36	30	146.3	162.5	129.3	14.5	0	0	0	1	128	2

churn

Valmis70%

CLASSICAL

The new deal with data.



NETWORK / CONNECTED APPROACH

# (sotsiaal)**võrgustike** analüüs?

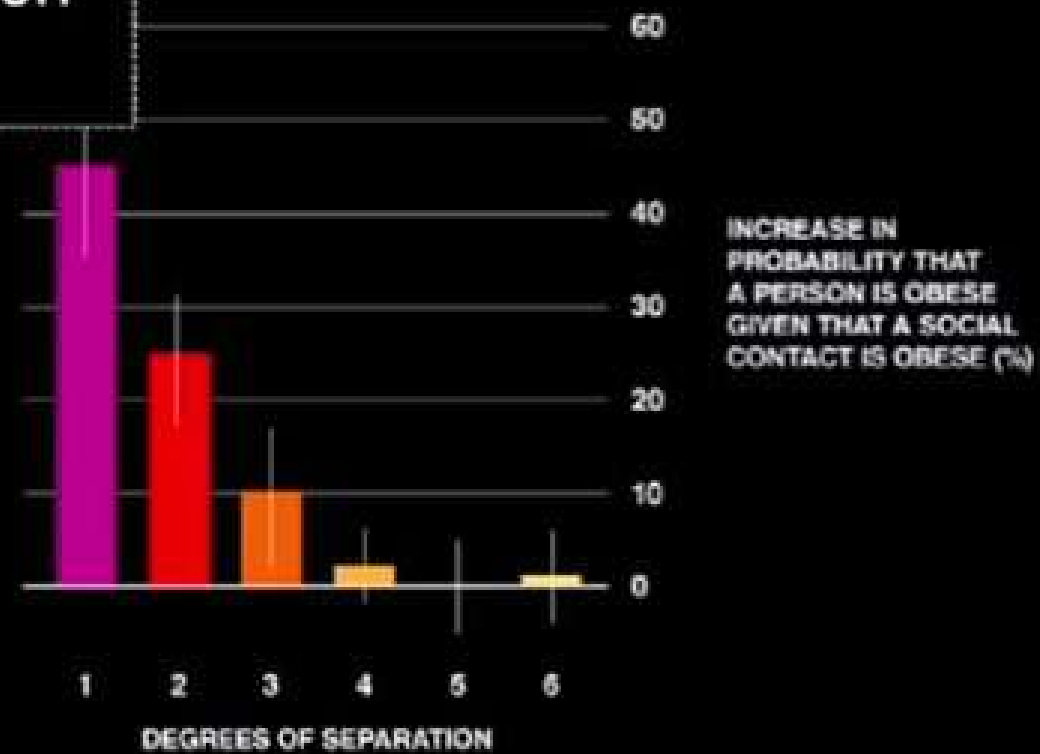
- Mõtteviis ja komplekt tööriistad, et osapoolte omavahelist suhtlemist kaardistada ja analüüsida (SUUR PILT);
- **Sotsiaalvõrgustike analüüs ei ole Facebook'i analüüs**, vaid peetakse tänapäeval organisatsioonikäitumise alamteemaks ning omab juuri Moreno sotsiomeetria koolkonnas;
- Rahvusvaheliselt kasutusel andmebaasiturunduses (perepeale ja seltskonna liidritele on efektiivsem turundada!).



# Teoreetilised alused

- Milgrami eksperiment: kõik inimesed on teineteisega seotud kõige rohkem 6 inimese kaudu;
- Võrgustike liikmed omandavad võrgustikust, selle ülesehitusest ja oma paiknemisest uusi struktuurseid omadusi;
- Üldfilosoofiast tuntud: me ise kujundame oma võrgustikke ja võrgustikud kujundavad meid;
- Võrkudel on väärtus iseeneses - sotsiaalne kapital, pluss hulk omadusi kerkib esile just võrkude kindla struktuuri (ülesehituse) tõttu.

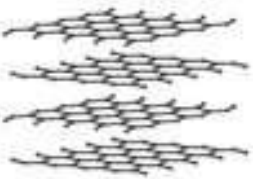
## Degree of Separation and Body Size



NÄIDE

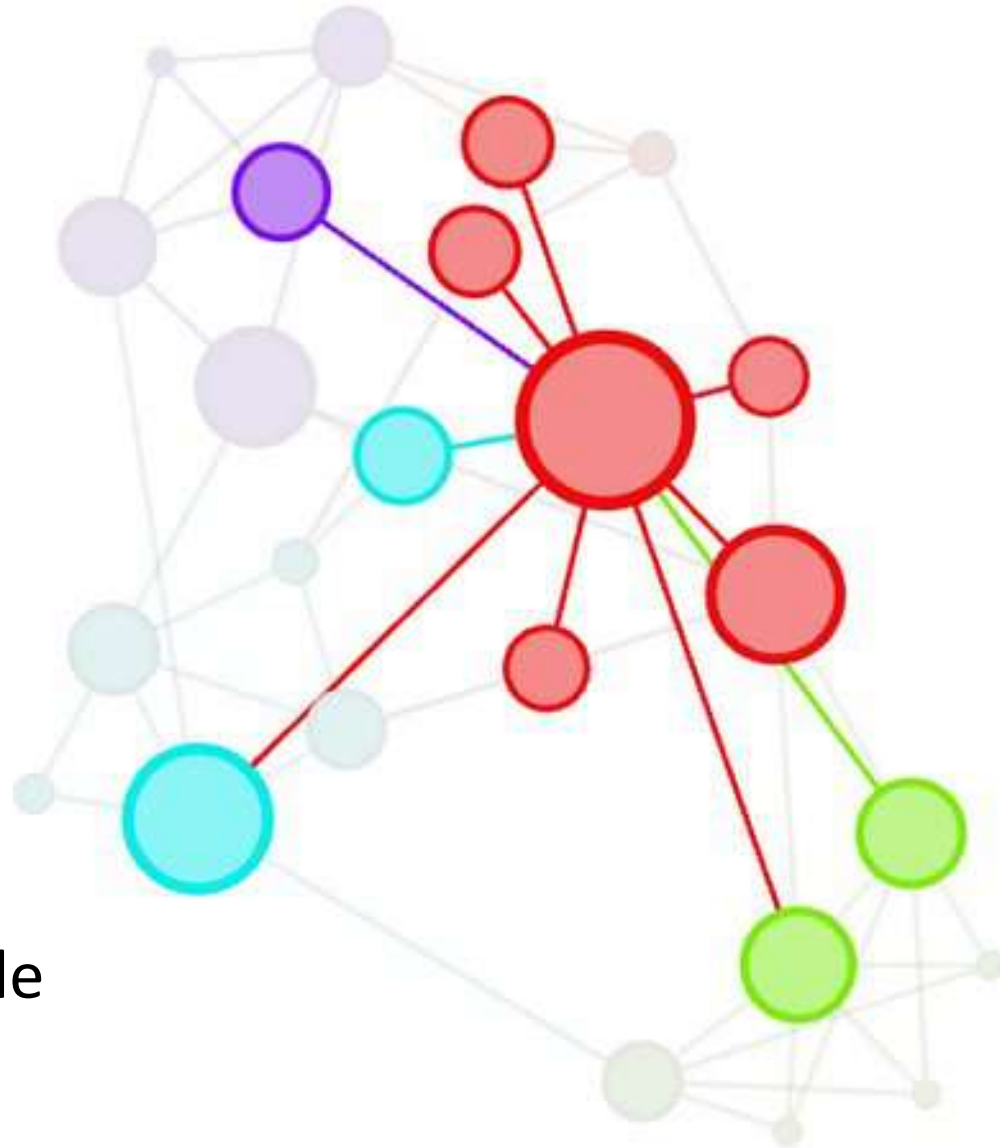
# Teemant ja grafiit (“harilik” pliiats)

- Hulk omadusi kerkib esile just võrkude kindla struktuuri (ülesehituse) tõttu;
- Mõlemad koosnevad süsiniku aatomitest, mis aga on “ühendatud” teisiti.
- Eri ühendus toob endaga teised omadused.



# Teoreetilised alused

- Tsentraalsusmõõdud (*centrality measures*)
- Suhete arv (*degree*)
- Lähedus kõigile (*closeness*)
- Vahelolek (*betweenness*)
- Kolm mõõtu on reeglina omavahel üsna seotud (korrelatsioonis)
- Mõtlemiseks: kas Te oskate oma kogemusest tuua näiteid, kus ei ole seotud?

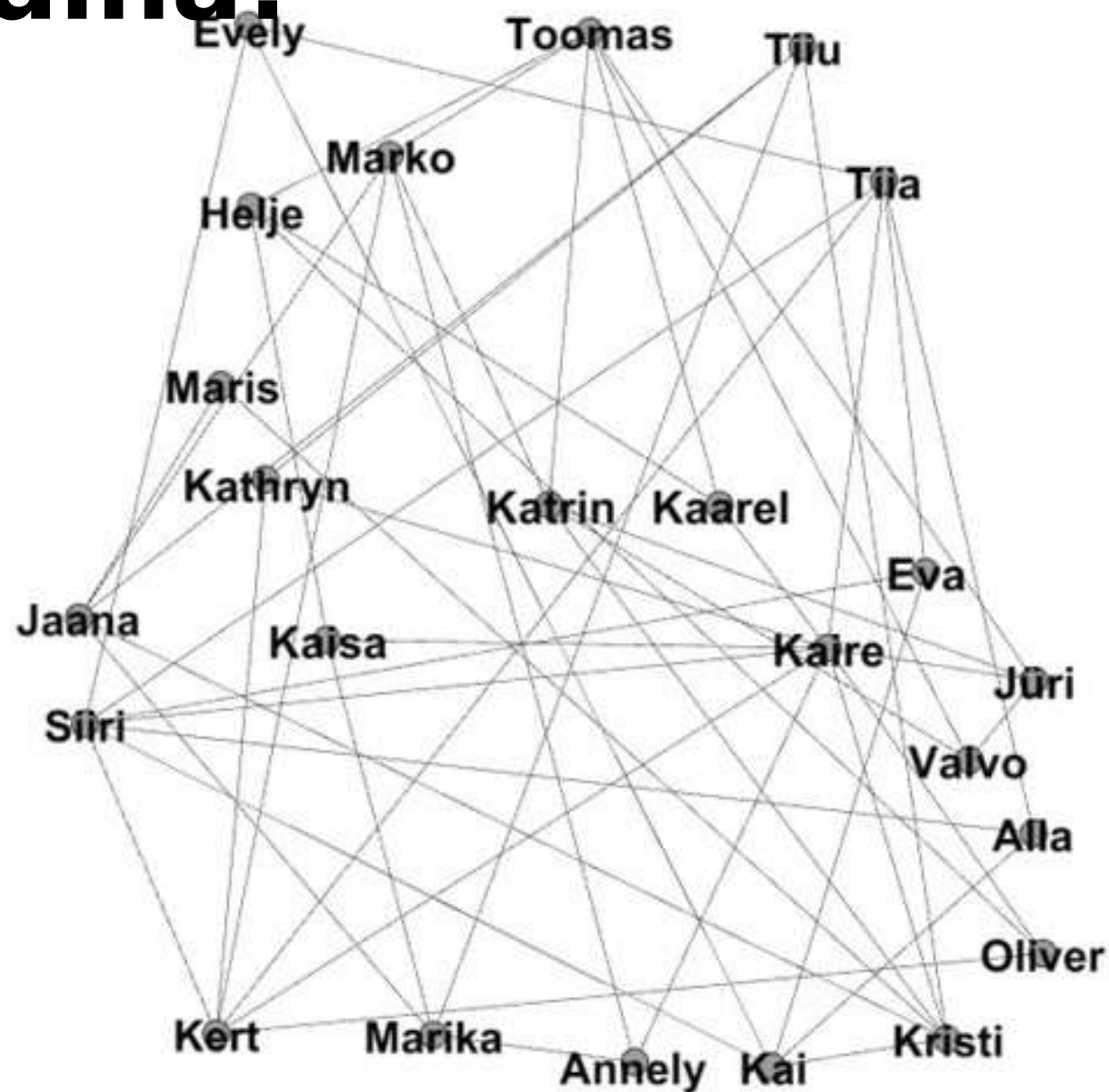




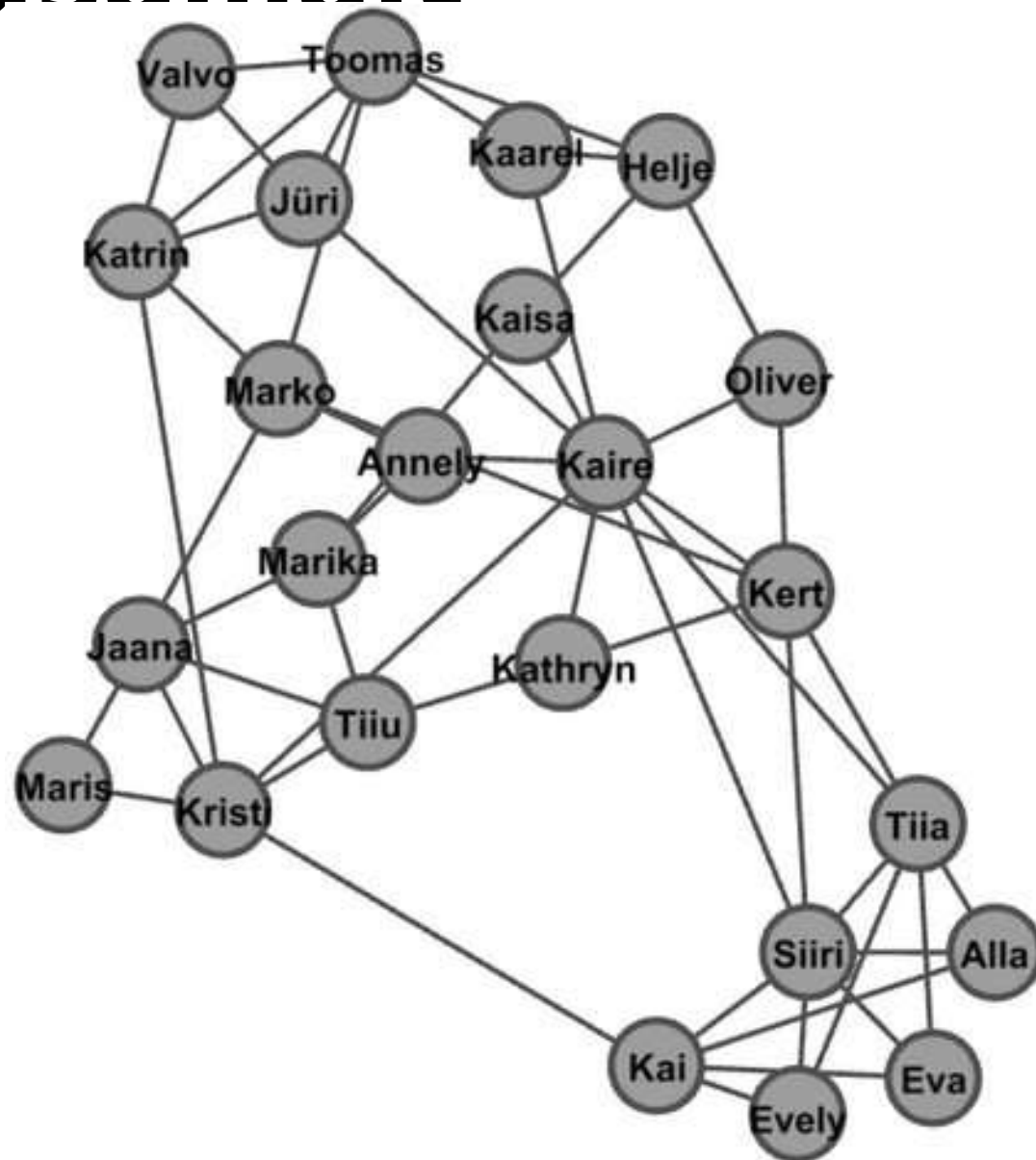
# **Kust alustada? Jälgige (küsi), kellega igaüks suhtleb kõige rohkem**

- Marko suhtleb kõige rohkem Kerdi, Jaana ja Katriniga. Jaana suhtleb kõige rohkem Kristi, Marise ja Markoga. Kristi suhtleb kõige rohkem Kai, Tiiu ja Katriniga. Maris suhtleb kõige rohkem Jaana ja Kristiga. Annely suhtleb kõige rohkem Kaire, Marika ja Markoga. Marika suhtleb kõige rohkem Kaisa, Jaana ja Tiiuga. Tiiu suhtleb kõige rohkem Kathrynini, Marika ja Jaanaga.
- Oliver suhtleb kõige rohkem Helje, Kaire ja Kerdiga. Toomas suhtleb kõige rohkem Kaarli, Marko ja Jüriga. Jüri suhtleb kõige rohkem Kaire, Toomase ja Valvoga. Kaarel suhtleb kõige rohkem Toomase, Kaire ja Heljega. Katrin suhtleb kõige rohkem Jüri, Toomase ja Valvoga. Valvo suhtleb kõige rohkem Jüri, Toomase ja Katriniga. Kert suhtleb kõige rohkem Siiri ja Kairega. Siiri suhtleb kõige rohkem Kerdi, Tiia ja Kairega. Evely suhtleb kõige rohkem Siiri, Tiia ja Kaiga. Eva suhtleb kõige rohkem Kai, Siiri ja Tiiaga. Alla suhtleb kõige rohkem Siiri, Tiia ja Kaiga. Kai suhtleb kõige rohkem Eva, Siiri ja Allaga. Tiia suhtleb kõige rohkem Siiri, Kerdi ja Kairega. Kathryn suhtleb kõige rohkem Kert, Kaire ja Tiiuga. Helje suhtleb kõige rohkem Kaisa, Oliveri ja Toomasega. Kaisa suhtleb kõige rohkem Kaire, Helje ja Marikaga. Kaire suhtleb kõige rohkem Kristi, Kerdi ja Siiriga.

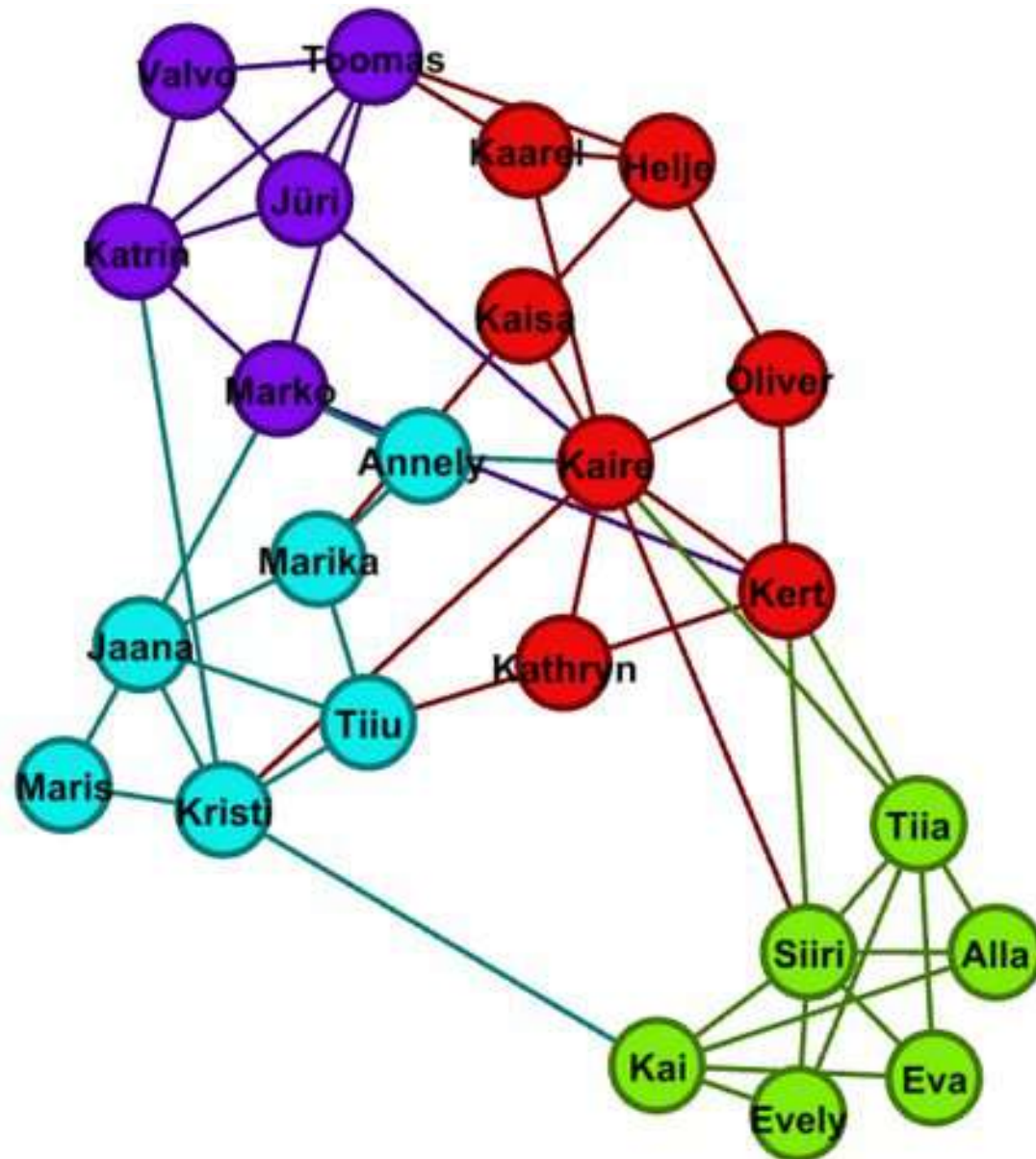
# Süstematiseerige info pildina!



# Joonistame sama asja “ilusamalt”

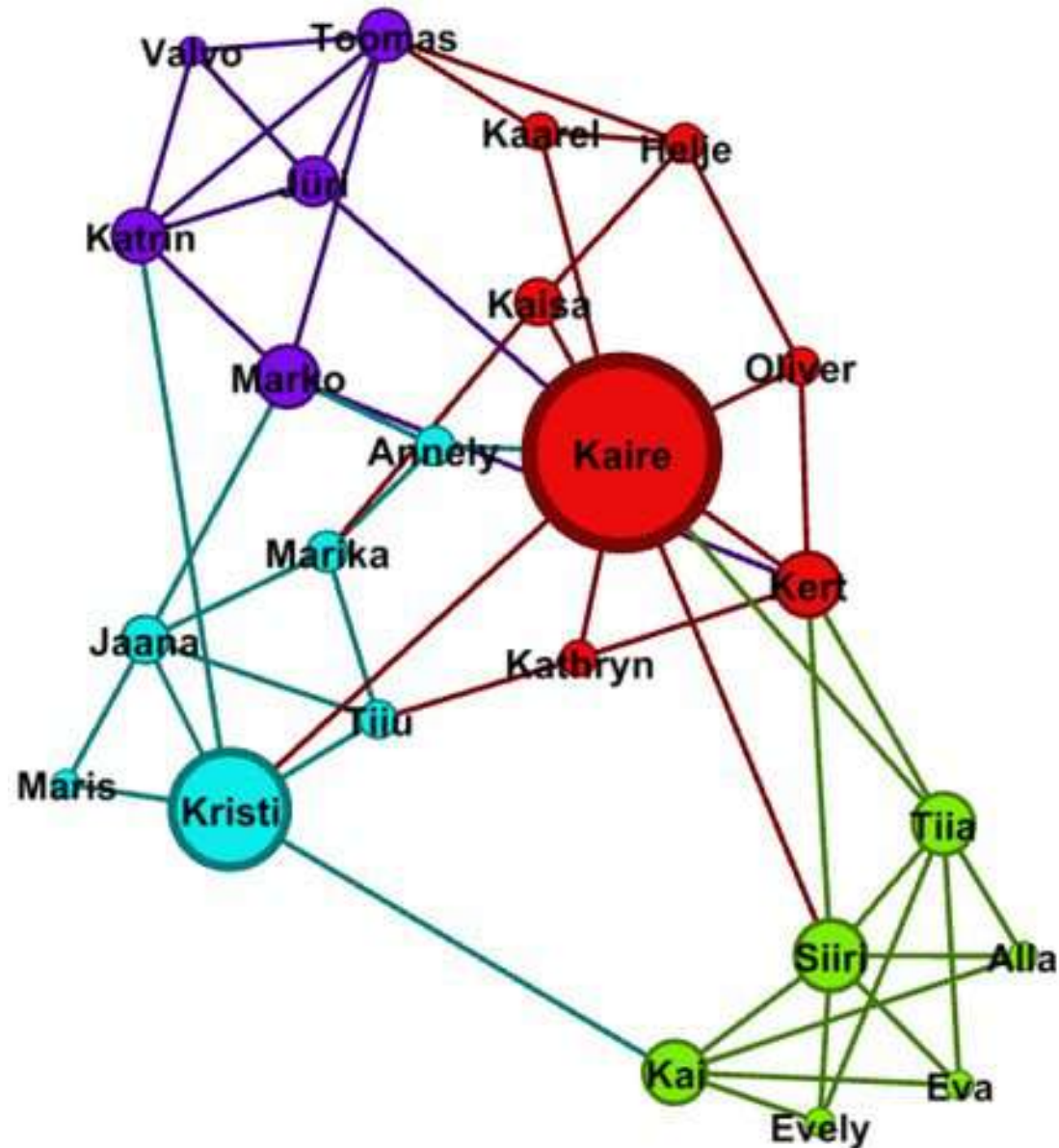


# Leiame sõpruskonnad

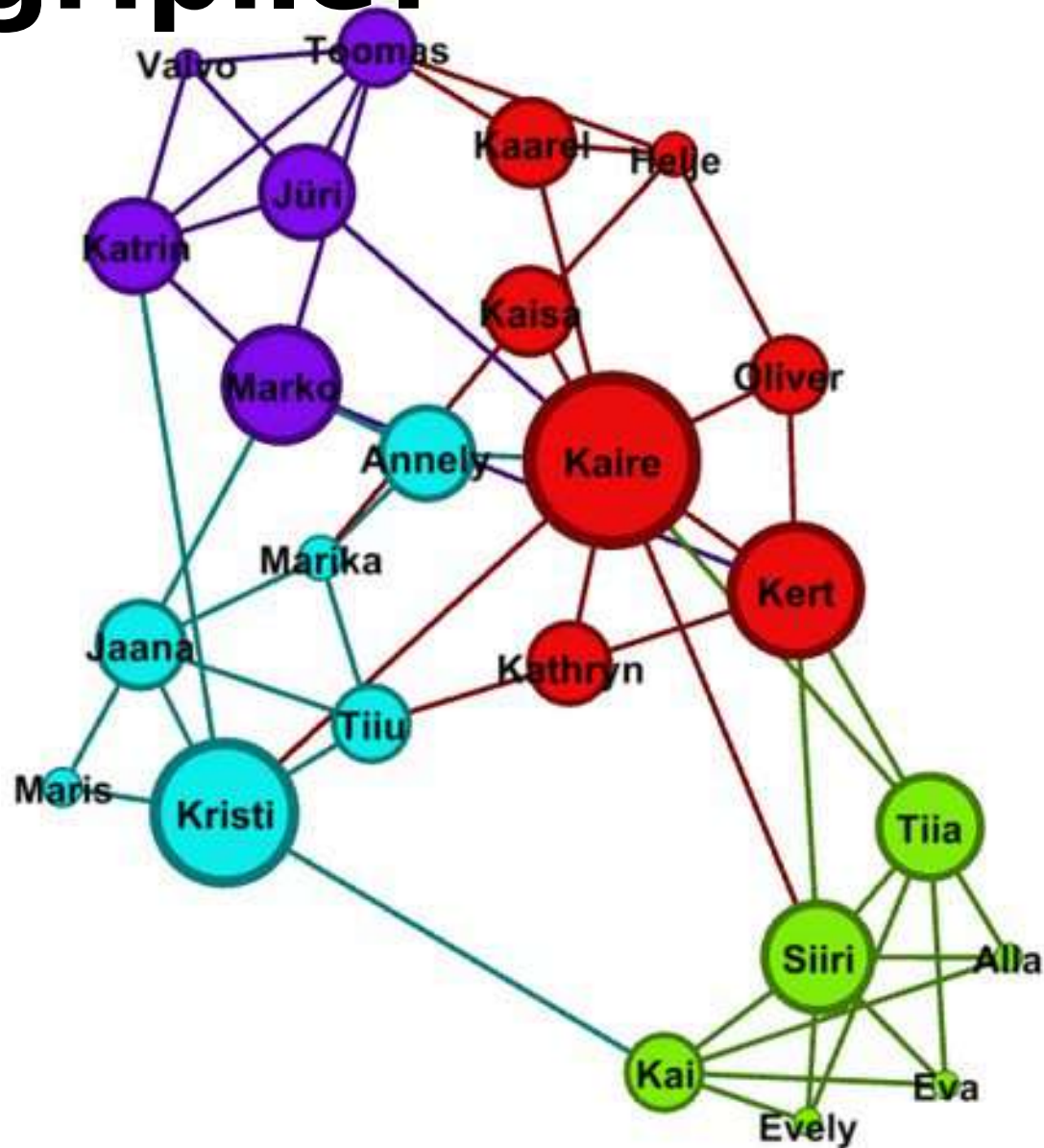




# Leiame parimad “vahendaiad”

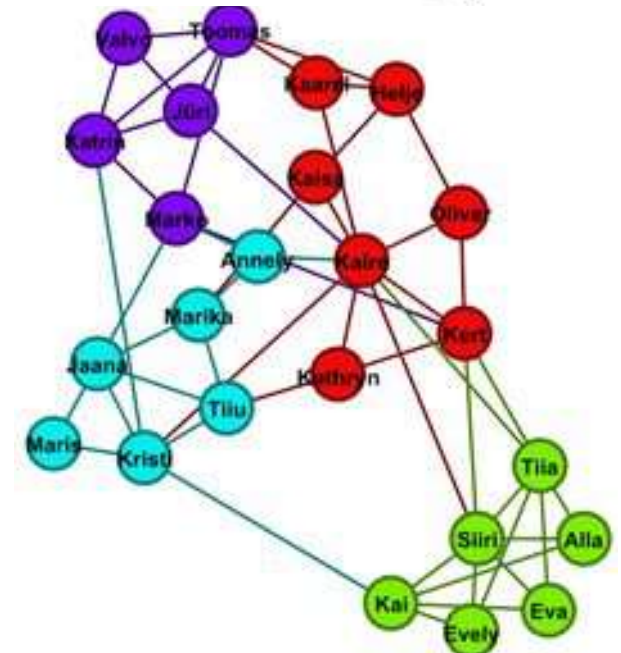
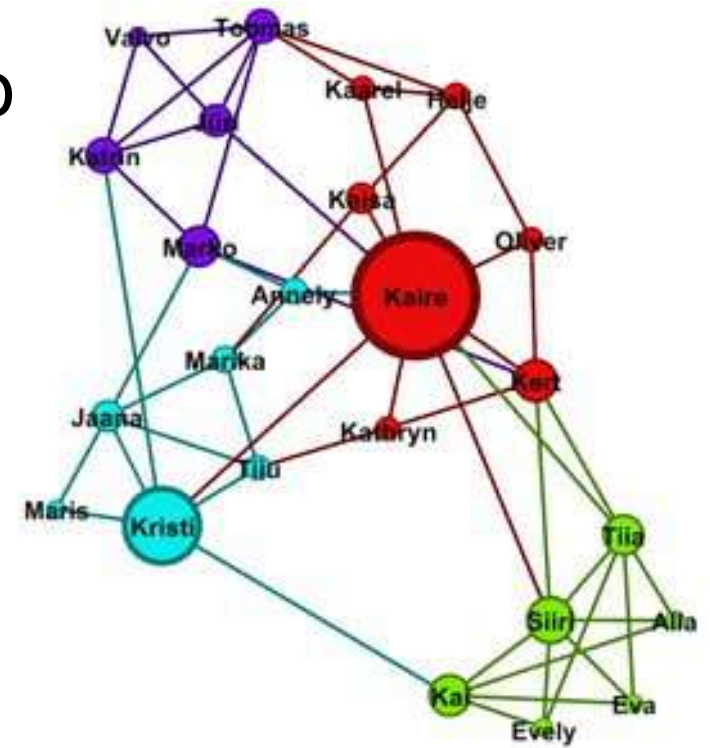


# Kellel on ligipääs infole ja gripile?



Suuruse, värvi ning seose **paksuse** abil saab kõike kirjeldada

- Rahulolu
- Kasumlikkus
- Erinevad kompetentsid
- Kompetentsid (näiteks erinev kompetents eri värviga) ning tulemused (näiteks suurus!)



# Mõtlemiseks

- Mis on mulle olulised uued andmeallikad?
- Kuidas ma saaksin neid (koos) kasutada?
- Mida peaks selle infoga peale hakkama?  
Keda huvitab? Millised otsused vajavad  
kiiremat infot uudsetest allikatest?  
Vajame me kiiremat informatsiooni?



“If you torture the data  
long enough, it will  
confess.”

— Ronald H. Coase, *Essays on Economics and Economists*

